DOCUMENT RESUME

ED 160 663 TM 007 971

AUTHOR Stetz, Frank P.; Beck, Michael D.
TITLE A Survey of Opinions Concerning Users of Educational Tests.
PUB DATE Mar 78
NOTE 26p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Toronto, Ontario, Canada, March 27-31, 1978)

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS Administrator Attitudes; Communication Problems; *Educational Testing; Group Membership; Knowledge Level; *National Surveys; *Opinions; Professional Associations; *School Superintendents; *Specialists; Testing Problems; Test Interpretation; *Use Studies
IDENTIFIERS *National Council on Measurement in Education

ABSTRACT

Questionnaires concerning the users of educational tests were returned by 87 school superintendents in the United States and 228 members of the National Council on Measurement in Education (NCME). Most NCME respondents were associated with colleges. The NCME sample represented measurement experts, while the superintendents represented non-measurement-expert users of tests--persons who do not have a degree related to this field of specialization. Both groups were surveyed regarding: (1) measurement issues of greatest concern to the non-expert users; (2) ways to improve communication between the two groups; (3) measurement topics of interest to non-experts which would stimulate their interest in NCME; (4) necessary level of measurement proficiency for effective functioning of non-experts; and (5) testing trends of importance to test users. There was a major discrepancy between each group's opinion of user needs toward intelligence testing, norm and criterion-referenced testing, and the use of standardized tests for accountability. NCME members' opinions of user needs were closer to those of testing critics, while superintendents were far more positive. The groups did agree on the subject of the grade level in which test users are primarily interested, and on purposes of testing, necessary proficiency levels and measurement topics of interest. The gap between the groups is wide enough to raise the question of whose perceptions are closest to those of the total population of nonmeasurement-expert users of tests. (Author/GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

FRANK P. Stetz
Michael D Beck

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

ED160663

A Survey of Opinions Concerning Users of Educational Tests

Frank P. Stetz & Michael D. Beck

The Psychological Corporation
New York City

Paper presented at the annual meeting of the National Council on Measurement in Education, Toronto, Ontario, March 1978

Printed in the United States of America

TM 007 971

A Survey of Opinions Concerning Users of Educational Tests

Frank P. Stetz & Michael D. Beck
The Psychological Corporation

The purposes of this study were to determine: 1) measurement topics of greatest concern to "non-measurement-expert" users of educational tests; 2) ways to improve communications between NCME members and the "non-measurement-expert" group; 3) opinions on the measurement topics of interest among "non-expert" users which would stimulate their interest in NCME; 4) levels of measurement proficiency necessary for "non-measurement-expert" users to function effectively; and 5) testing trends seen as important to such users. For the purposes of this survey, a "non-measurement-expert" user was defined as a person directly involved in education tests and measurements who does not have a degree related to this field of specialization. Random samples of NCME members and public school superintendents were polled.

TM007 971

Paper presented at the annual meeting of the National Council on Measurement in Education, Toronto, Ontario, March 1978

Introduction

In a symposium chaired by Stetz & Beck (1977) at the 1977 annual meeting of NCME, various perspectives of the organization were reviewed by participants. Irv Lehmann presented the history of the organization; Beck & Stetz analyzed the past ten annual meetings in terms of selected variables; Annie Ward discussed the interface of NCME with the measurement community; and David Payne and Bob McMorris contributed data on the topic of future goals and directions of the organization.

Ebel (1977), discussant for this symposium, stressed some points concerning users of tests in his summation of the program, including:

(1) Users of tests are not as prominent in NCME membership as we would like, but perhaps as prominent as they care to be.

(2) NCME members see needs in education that better measurement could help to meet, if only they were used more widely and effectively.

(3) NCME members would like to see the organized efforts of NCME be applied to bring about that wider and more effective use.

Purposes

The common thred in the above points is the need to examine the status of NCME in relation to this larger group of "non-measurement-expert" users of tests. For the purposes of this survey, "non-measurement-expert" users are defined as those directly involved with educational tests and measurements who do not have degree related to this field of specialization. The purpose

of this study was to survey the opinions of a random sample of NCME members concerning such topics as:

- The status of "non-measurement-expert" users within the NCME organization;
- Ways to improve communication between NCME members and "non-measurement-experts" users;
- Measurement topics perceived by NCME members to be of interest among "non-expert" users which would stimulate their interest in NCME;
- Levels of measurement proficiency necessary for users to function effectively; and
- Testing trends seen as important to such users.

The opinions of a random sample of school superintendents were also assessed.

The rationale for conducting the study was based in part on the increased criticism of educational test usage among various educational groups. Possibly this criticism stems in part from a lack of knowledge about tests. Those given the charge of working with test results may not be adequately informed of the advantages and limitations of test usage. As part of an ongoing dialogue between measurement experts and "non-measurement-expert" users of tests, an assessment of the understanding of the topic by both groups is needed. As a preliminary step in addressing this vast group of "non-measurement-expert" users, the authors "pilot tested" items on a group of measurement experts (NCME members) and school administration experts (superintendents)--two groups who come in direct contact with "non-measurement-expert" users. Results of this study can then be used to refine the types of questions asked directly of users.

Survey Procedures

Two forms of a questionnaire were developed. The items concerning"non-measurement-expert"users were identical on the two questionnaires; slightly different demographic data were collected for the two groups.

One part of the sample consisted of a 20 percent random sampling (N = 450) of NCME members provided by that organization. The other sample consisted of 450 superintendents of randomly selected school districts in the United States. The random list of schools was chosen to be representative of various geographic areas and school district enrollments. Each questionnaire was addressed to the attention of the superintendent. No persons were included in both samples.

Materials sent to each member of the sample consisted of an introductory letter, a questionnaire, a self-addressed stamped envelope and an incentive. The introductory letter stated the purpose of the questionnaire, defined the term "non-measurement-expert" user, explained the respective sample groups' relationships with this "non-expert" group, and stressed the importance of assisting with the study. In addition, the letter included instructions for the completion and return of the questionnaire.

For the NCME group, the incentive for cooperation was a "Testing -- Do Not Disturb" sign to "offer a few minutes respite from (their) busy schedules to answer our questions." The superintendents were offered a free test specimen set from a list that was provided. All respondents were offered a complimentary copy of the final report of this survey.

All questionnaires were mailed on October 1, 1977; the due date for return was November 15. On November 1st, a follow-up letter was sent to each member of the sample, reminding them of the due date and the importance of the study.

Results and Discussion

By the cut-off date, 228 NCME questionnaires and 87 superintendent questionnaires were returned. This constituted a 50.7% and a 19.3% response rate, respectively. Of the NCME respondants, 23% were females; females made up 18% of the respondents to the superintendent questionnaire. This 18% figure appears to be quite high, given the number of female school superintendents in the country. Some superintendendents may have passed on the questionnaire to a female member of their staff for completion.

The majority of the NCME respondents were located in college departments of education, psychology, educational psychology, or educational measurement. The second most frequent institutional affiliation was the category "city, county or state department of education." NCME respondents cited teaching, program evaluation, research, and administration most frequently as their primary job function.

In general, respondents from the superintendent group were divided evenly between school districts enrolling "under 1000" or "1000 to 10,000" students. Only seven questionnaires were returned from large school districts (enrollment over 10,000). A majority of the superintendents characterized their systems as rural; one-quarter described their system as suburban; approximately 10% were urban systems.

Geographically, the respondents were spread rather evenly across the country. Figure 1 outlines the geographic distribution of all respondents according to the first zip code digit of their address.

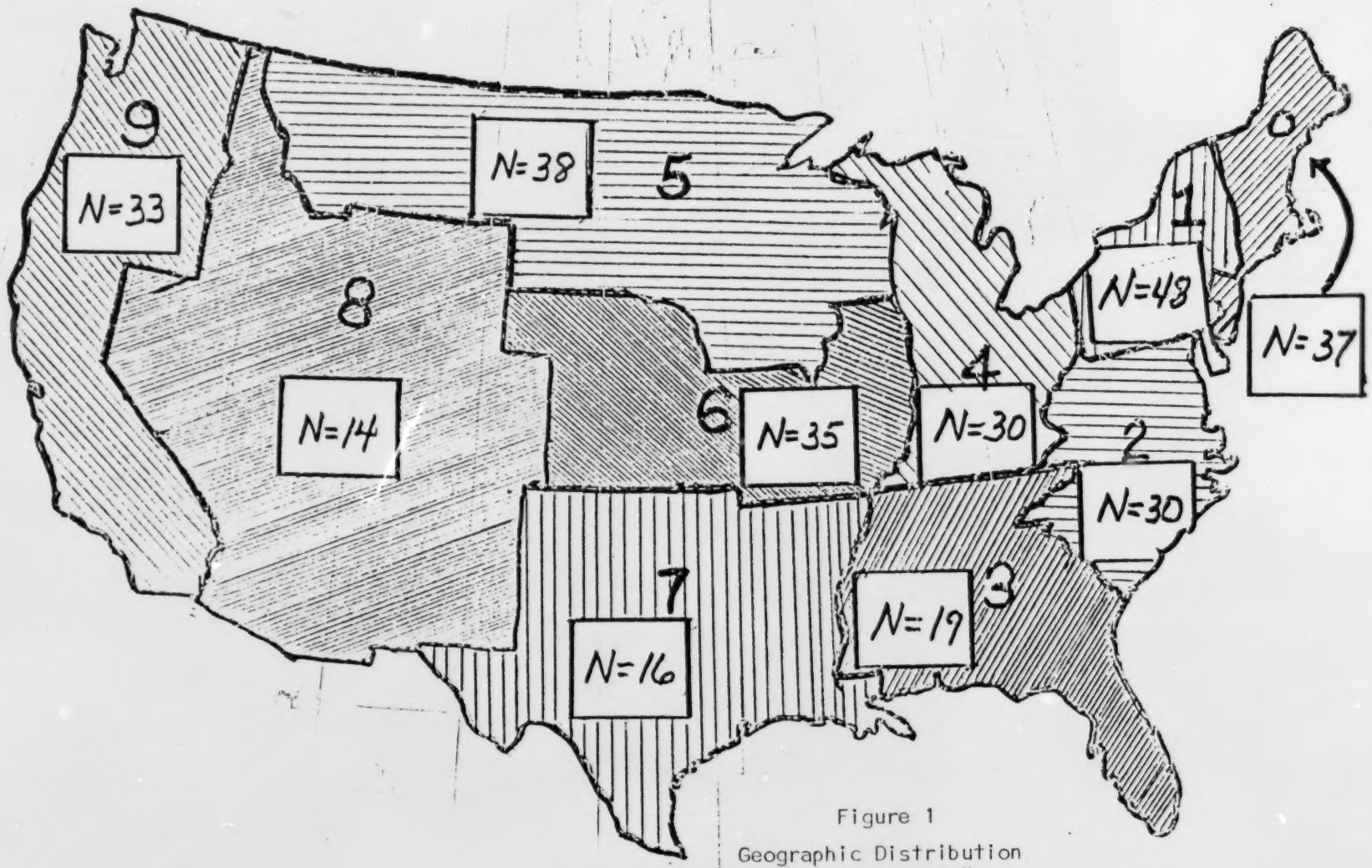

Figure 1

Geographic Distribution
of Respondents*

*There were, in addition, nine respondents with foreign addresses.

THE NCME and superintendent sample respondents averaged 2.6 and 1.6 professional organization memberships respectively. (Membership in NCME was not included in the tabulations.) Male NCME respondents averaged 2.5 memberships while the females' average was 2.3. On the other hand, females in the superintendents' sample were, on the average, members of 2.1 organizations; the male average was 1.5.

Tables 1, 2, and 3 summarize the respondents' opinions of the areas of testing, environments, and major measurement issues with which "non-measurement-expert" users are most concerned. A majority of NCME respondents felt that current achievement status, selection, mental ability, and mental processes were the primary areas of concern to such users. Superintendents' responses were similar, but they rated users as being more concerned with diagnosis than selection. They also were not as concerned with specific mental processes. Diagnosis and selection were the two areas in which the responses of the two groups differed most. Responses were similar for both groups on the subject of environments with which users are most concerned. As anticipated, elementary, secondary, and preschool education were judged to be of primary interest to users.

Superintendents felt that the "non-measurement-expert" users were most concerned with the reporting and interpretation of test results and the measurement of growth. NCME members polled agreed that test interpretation was a major concern, but felt that users were more concerned with cultural and sex bias than with the measurement of growth. The two groups differed most in their opinions on the bias issue.

Table 4 summarizes the results from four semantic differential scales designed to assess respondents' interpretations of users' mood toward intellgience testing, norm- and criterion-referenced testing, and the use of standardized

Table 1

Respondents' Opinions of the Areas of Testing and Test Use With Which Users Are Most Concerned

	NCME Members' Views	Superintendents' Views
	%	%
General mental ability or scholastic aptitude	18	30
Specific mental processes such as problem solving	18	3
Diagnosis	3	24
Selection	19	2
Interest measurement	4	7
Current achievement status	28	29
Psychomotor skills	1	2
Personality/attitude measurement	9	3

Table 2

Respondents' Opinions of the Envirnoments With Which Users are Most Concerned

	NCME Members' Views	Superintendents' Views
	%	%
Preschool/early childhood education	14	15
Elementary education	28	33
Secondary education	25	30
Vocational/job training	10	13
Junior/community college	2	1
Undergraduate college or university	9	4
Teacher education (graduate or undergraduate)	4	1
Other professional education	4	-
Adult basic and continuing education	4	3

Table 3

Respondents' Opinions of the Major Measurement Concerns and Issues With Which Users Are Most Concerned

	NCME Members' Views	Superintendents' Views
	%	%
Item writing/test construction techniques	9	3
Test format and administration considerations	6	6
Alternate scoring practices and techniques	1	1
Procedures and techniques for reporting test results	14	11
Issues of cultural or sex bias in test construction/interpretation	12	2
Concerns in the measurement of bilingual examinees	1	2
Ethical and legal issues in measurement	3	1
Proper interpretation of test results	11	18
Sampling and research design methodology	1	1
Validity studies	2	3
Measurement of "growth"	8	13
Criterion-referenced measurement	8	9
Performance testing/competency-based evaluation	6	9
Evaluation of curriculum instructional materials	3	6
Evaluation of programs, title projects, etc.	8	7
Evaluation/certification/licensing of teachers, other professionals	2	1
Minimal competency	5	7

Table 4

Respondents' Views of the Mood of "Non-Measurement-Expert" Users
Toward Various Testing Issues

Dimensions*	INTELLIGENCE TESTING		STANDARDIZED, NORM-REFERENCED ACHIEVEMENT TESTING		CRITERION-REFERENCED TESTING		USE OF STANDARDIZED TESTS FOR ACCOUNTABILITY PURPOSES	
	NCME Members	Sup'ts	NCME Members	Sup'ts	NCME Members	Sup'ts	NCME Members	Sup'ts
	%	%	%	%	%	%	%	%
Supportive								
1-2	17	33	30	51	45	50	8	12
3-5	63	62	63	49	54	46	41	55
6-7	20	5	7	0	1	4	51	33
Antagonistic								
Calm								
1-2	9	24	16	37	20	26	5	13
3-5	64	69	68	57	70	60	35	41
6-7	27	7	16	6	10	14	60	46
Anxious								
Interested								
1-2	32	45	39	55	47	51	27	32
3-5	63	52	58	44	52	49	54	59
6-7	5	3	3	1	1	0	19	9
Uninterested								
Knowledgable								
1-2	10	16	12	27	10	16	12	16
3-5	49	67	69	67	52	63	53	63
6-7	41	17	19	6	38	21	35	21
Not Knowledgable								
Sample Size	206-223	80-87	215-222	80-83	214-221	80-82	221-223	81-84

*Although the questionnaire used a seven-point scale, data were collapsed into three categories for summary purposes.

tests for accountability purposes. The semantic differentials were 7-point scales on the questionnaire, but for interpretive ease, these were collapsed into three intervals in Table 4. NCME respondents thought that users were, for the most part, supportive toward standardized, norm-referenced achievement testing and criterion-referenced testing, ather neutral toward intelligence testing, and antagonistic toward the use of standardized tests for accountability purposes. The superintendents who responded felt that users were more supportive toward all four issues than did the NCME sample. The largest difference between the two samples was on the issue of standardized, norm-referenced achievement testing. Thirty percent of the NCME sample and 51% of the superintendents felt that users were supportive.

Many more NCME respondents than superintendents felt that users were anxious about testing; the superintendents who responded consistently felt that users were more calm about tests. Once again, the greatest difference in opinion between the two groups showed up in rating users' moods toward standardized, norm-referenced testing.

The "interest" continuum showed the greatest similarity in opinions for the two groups. Both groups felt that users in general were interested in all four of the testing issues. Again, however, across all four questions, the superintendent group rated users as more interested in the issues than did NCME members.

The "knowledgable/not knowledgable" scale showed the greatest discrepancy in opinion in the area of intelligence testing. NCME respondents rated users as less knowledgable about the testing issues listed than did the superintendents.

Table 5 summarizes respondents' views of the necessary measurement proficiency levels required for teachers, principals, guidance counselors, and test directors. Some respondents had difficulty in responding to this question; consequently the sample sizes for this question were reduced. Nevertheless, the majority of respondents in both groups answered similarly. That is, the plurality of both groups felt that teachers, principals, and guidance counselors should be required to pass a comprehensive test and measurements course before being allowed to work with these instruments in schools. The plurality in both groups also felt that test directors should be licensed. The question remains, licensed by whom?

Members of both samples were asked which areas they felt "non-measurement-expert" users would be most interested in learning more about. Table 6 summarizes their responses. Criterion-referenced testing and interpretation, minimal competency, and accountability were the leading choices of both samples. In addition, diagnostic testing was mentioned frequently by the NCME respondents. None of the other topics were rated as being of interest by over 8% of either sample.

When asked which means could be used to improve communications with "non-measurement-expert" users, many NCME respondents felt that 1) a coordinated effort among measurement associations, 2) measurement association-sponsored, low cost travelling workshops/seminars, and 3) measurement associations lobbying for mandatory tests and measurement courses for all teacher education students would improve communication (see Table 7). The most popular choice among superintendents was the second point above; many also felt that improved communications could be achieved through the creation of a journal designed specifically for "non-measurement-expert" users and measurement association-sponsored sessions at educational association meetings.

Table 5

Respondents' Views of Necessary Measurement Proficiency Levels for Various School Personnel[1,2]

Response Category	NCME Members' Views				Superintendents' Views			
	Teacher	Principal	Guidance Counselor	Test Director	Teacher	Principal	Guidance Counselor	Test Director
All people involved in tests and measurements should be licensed.	-	-	17	42	5	3	16	32
All people involved in tests and measurements should be required to pass a national proficiency examination before being allowed to work with tests and measurements	5	11	30	18	-	4	9	12
All people involved in tests and measurements should be required to PASS a comprehensive tests and measurements course before being allowed to work with tests and measurements.	50	48	36	30	35	30	62	28
All people involved in tests and measurements should be required to ENROLL IN a comprehensive tests and measurements course before being allowed to work with tests and measurements.	10	10	9	4	16	26	2	5
All people involved in tests and measurements should be required to attend inservice seminars in tests and measurements.	25	22	3	1	28	28	2	9
No regulatory steps should be mandated for people involved with tests and measurements.	10	9	5	5	16	9	9	14

[1] Numbers are expressed in percents.

[2] Due to the difficulty some respondents experienced in answering this question, sample sizes were reduced to 186 (NCME) and 43 (superintendents).

Table 6

Respondents Perceptions of Further Interest of "Non-Measurement-Expert" Users in Certain Educational Testing Topics

	NCME	Superintendents
	%	%
Behavioral objectives for each test item	4	8
Computer-assisted testing	6	4
Norm-referenced testing and interpretation	3	4
Criterion-referenced testing and interpretation	21	23
Matrix sampling testing	4	4
State assessment	3	4
Minimal competency	17	12
Accountability	11	14
Diagnostic testing	12	4
Intelligence testing	1	4
Assessment of non-cognitive variables (e.g., attitudes)	7	6
Use of test results to determine promotion/graduation	4	6
Use of test results for teacher evaluation	5	4
Use of test results for college/professional school admission	2	3

Table 7

Respondents' Opinions of Possible Activities to Improve Communication With Users

Activity	NCME Members' Views	Superintendents' Views
	%	%
The creation of a "non-measurement-expert" user professional organization	3	2
The creation of a journal designed specifically for non-measurement-expert users	11	17
A coordinated effort among measurement associations (for example, the National Council on Measurement in Education and the Association for Measurement and Evaluation in Guidance) in reaching and assisting non-measurement-expert users of educational tests	21	14
Measurement associations lobbying for mandatory tests and measurement courses for all teacher education students	13	7
Encourage measurement association members to publish in non-measurement-expert users educational journals	11	8
Identify experts in educational measurement who work in local educational authorities and departments of education and use them to develop materials, etc.	10	9
Measurement association-sponsored low-cost travelling workshops/seminars for non-measurement-expert users of educational tests.	20	24
Measurement association-sponsored sessions at educational association meetings.	11	19

Table 8 summarizes respondents' predictions of the status of certain educational testing concepts ten years from now. If such predictions come true, these will be topics with which users must deal. In general, the responses of both groups were quite similar. Many in both groups felt that state assessment, minimal competency testing, and accountability will be mandated in ten years. A majority of both groups felt that all the topics listed would have increased acceptance or be mandated by then with the exception of intelligence testing. Many in both groups though there would be no change from the present for norm-referenced testing, matrix sampling testing, intelligence testing and the use of test results for college/professional school admission. Interestingly, NCME respondents were equally split on whether there would be no change from the present or less attention paid to intelligence tests; there was a similar split for the superintendent group.

Conclusions

While the two groups who come into direct contact with "non-measurement-expert" test users agree on many points concerning such users, there is a wide enough gap between the two groups to raise the issue of whose perceptions are closest to those of this large user group. With a few exceptions, the similarities in perceptions were most apparent in the issues dealing with environments, testing issues, proficiency levels necessary, and interests. The issue of moods of users toward testing shows the major discrepancies. NCME members' opinions of user moods are closer to those of testing critics. Superintendents, in general, tended to be far more positive.

Implications for Further Research

The next step in the process of improved communications and improved understanding of this large body of test users would be to poll their feelings and opinions. It would be interesting to see whether their general mood is closer in truth to the views of the sampled NCME members or to those of the superintendents polled.

TABLE 8

Respondents Predictions of
the Status of Certain Educational Testing Topics Ten Years from Now

Topic	Mandated by		Increased Acceptance		No Change from Present		Less Attention Paid		Outdated/Forgotten	
	NCME %	Sup't %	NCME %	Sup't %	NCME %	Sup't %	NCME %	Sup't %	NCME %	Sup't %
Behavioral objectives for each test item	7	18	54	53	18	14	15	9	6	6
Computer-assisted testing	1	11	81	71	13	13	4	4	1	1
Norm-referenced testing and interpretation	4	6	24	35	47	34	24	24	1	1
Criterion-referenced testing and interpretation	5	12	70	72	10	14	13	-	2	2
Matrix sampling testing -	2	7	52	33	38	43	6	11	2	6
State assessment +	33	55	43	27	18	11	5	5	1	2
Minimal competency +	44	60	41	31	9	7	5	-	1	2
Accountability	29	39	36	47	23	7	10	6	2	1
Diagnostic testing	11	31	67	52	19	15	3	1	-	1
Intelligence testing	1	9	9	16	42	30	42	37	6	8
Assessment of non-cognitive variables (e.g., attitudes)	1	5	53	58	36	21	9	12	1	4
Use of test results to determine prmotion/graduation +	14	25	47	47	27	18	10	8	2	2
Use of test results for teacher evaluation	1	2	39	53	44	28	12	11	4	6
Use of test results for college/professional school admission	7	12	21	37	57	42	13	8	2	1

Once opinions from this group of users are obtained, planning for an improved communications flow between the experts and the users can begin.

References

Ebel, R. L., Personal Communication, May 11, 1977

Stetz, F. P., & Beck, M.D. A review of selected perspectives of a national professional organization, or can NCME find true happiness after 40? Symposium presented at the annual meeting of the National Council on Measurement in Education, New York, New York, April 1977.